

True Feb 11 13:12:50 2003

us-09-831-426c-2.rsp

Page 1

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 10, 2003, 13:06:36 : Search time 21 Seconds

(without alignments)  
720.898 Million cell updates/sec

Title: US-09-831-426c-2

Perfect score: 2008

Sequence: 1 MDPFAVAVSVSLTRIADAF.....ESPNCVEDKMLSTVAVLTLG 365

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt-40.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1993	99.3	423	TF3A_HUMAN	O92664 homo sapien
2	1055.5	92.6	366	TF3A_XENLA	P03001 xenopus lae
3	1047	52.1	335	TF3A_RANPT	P36699 zena pipien
4	1024	51.0	339	TF3A_XENBO	P17842 xenopus bor
5	1022	50.9	329	TF3A_BURFM	P31694 bufu americ
6	816	40.6	322	TF3A_ICTPU	P79797 ictalurus p
7	556.5	27.7	365	P43_XENLA	P25455 xenopus lae
8	553.5	27.6	365	P43_XENBO	P25066 xenopus bor
9	417.5	20.8	803	2XDE_HUMAN	P98169 homo sapien
10	410.5	20.4	616	2G57_XENLA	P35789 homo sapien
11	409	20.4	336	2G57_XENLA	P18729 xenopus lae
12	409	20.2	799	2XDA_HUMAN	P98168 homo sapien
13	406	20.2	698	2Z34_HUMAN	O14588 homo sapien
14	404	20.1	821	2N41_HUMAN	O14588 homo sapien
15	403	20.1	574	VE73_HUMAN	O98225 homo sapien
16	399	19.9	1167	2Z08_HUMAN	O45343 homo sapien
17	398.5	19.8	626	2143_HUMAN	O45343 homo sapien
18	398.5	19.8	721	2337_HUMAN	O93744 homo sapien
19	395.5	19.7	433	206_XENLA	O93744 xenopus lae
20	395	19.7	803	2Z26_HUMAN	O94745 homo sapien
21	394.5	19.6	803	2N43_HUMAN	O94745 homo sapien
22	393	19.6	489	2Z1F_HUMAN	O94745 homo sapien
23	393	19.6	689	PRD9_HUMAN	O94745 homo sapien
24	392.5	19.5	337	2N81_HUMAN	P51508 homo sapien
25	391	19.5	738	2N84_HUMAN	P51523 homo sapien
26	386.5	19.2	759	1287_MOUSE	O96989 mus musculu
27	386.5	19.2	898	2Z07_XENLA	P18751 xenopus lae
28	385	19.2	913	2228_HUMAN	O94745 homo sapien
29	385	19.2	625	2189_HUMAN	O94745 homo sapien
30	383	19.0	538	2185_HUMAN	O12904 homo sapien
31	382	19.0	280	2G46_XENLA	P17722 xenopus lae
32	382	19.0	645	2F93_MOUSE	O61116 mus musculu
33	381.5	19.0	535	1	O94745 homo sapien

#### ALIGNMENTS

34	381.5	19.0	724	1	2N42_HUMAN	P28698 homo sapien
35	381	19.0	711	1	2175_HUMAN	O94743 homo sapien
36	380.5	18.9	595	1	2N85_HUMAN	O03923 homo sapien
37	380.5	18.9	682	1	2N45_HUMAN	O02386 homo sapien
38	379	18.9	683	1	Y798_HUMAN	O94745 homo sapien
39	378	18.8	652	1	Y798_HUMAN	O94745 homo sapien
40	377.5	18.8	595	1	2317_HUMAN	O96989 homo sapien
41	377	18.8	604	1	2300_HUMAN	O96989 homo sapien
42	376.5	18.8	446	1	2N70_HUMAN	O96989 homo sapien
43	376.5	18.8	472	1	21K3_HUMAN	O96989 homo sapien
44	376.5	18.8	570	1	2N76_HUMAN	P35508 homo sapien
45	376.5	18.8	697	1	HR1_HUMAN	P10072 homo sapien

RESULT 1

TF3A_HUMAN	STANDARD:	PRT:	423 AA.
AC	O92664: Q13097: Q12963:		
DT	01-NOV-1997 (Ref. 35, Created)		
DT	01-NOV-1997 (Ref. 35, Last sequence update)		
DT	16-OCT-2001 (Ref. 40, Last annotation update)		
DE	Transcription factor TF3A (Factor A) (TF3A).		
GN	TF3A.		
OS	Homo sapiens (human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.		
OX	NCBI_TaxID:9606;		
RM	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Fetal brain;		
RX	MEDLINE-9530928: PubMed-7789179;		
RA	Arakawa H., Nagase H., Hayashi N., Ogawa M., Nagata M.,		
RA	Fujiwara T., Takahashi E., Shii S., Nakamura Y.;		
RT	novel human gene (GTF3A) that is highly homologous to Xenopus		
RT	transcription factor TF3A.		
RL	Cytogenetic. Cell Genet. 70:235-238(1995).		
RM	[2]		
RP	SEQUENCE OF 61-423 FROM N.A.		
RC	TISSUE=Brain;		
RX	MEDLINE-95347600: PubMed-7622052;		
RA	Drew P.D., Nagle J.W., Canning R.D., Ozato K., Biddison W.E.,		
RA	Becker K.G.;		
RT	*Cloning and expression analysis of a human cDNA homologous to		
RT	Xenopus TF3A.		
RL	Gene 159:215-218(1995).		
RM	[3]		
RP	CHARACTERIZATION.		
RX	MEDLINE-94342241: PubMed-8063702;		
RA	Moorefield B., Roeder R.G.;		
RT	*Purification and characterization of human transcription factor		
RT	TF3A.		
RL	J. Biol. Chem. 269:20857-20865(1994).		
CC	-1- FUNCTION: INTERACTS WITH THE INTERNAL CONTROL REGION (ICR) OF		
CC	APPROXIMATELY 50 BASES WITHIN THE 5S RNA GENES, IS REQUIRED FOR		
CC	CORRECT TRANSCRIPTION OF THESE GENES BY RNA POLYMERASE III. ALSO		
CC	BINDS THE TRANSCRIBED 5S RNA'S. MAY INITIATE TRANSCRIPTION OF THE		
CC	5S RIBOSOMAL RNA GENE AND MAINTAIN THE STABILITY OF TRANSCRIPTION		
CC	OF OTHER GENES.		
CC	-1- SUBCELLULAR LOCATION: NUCLEAR.		
CC	-1- TISSUE SPECIFICITY: UBICUITOUS.		
CC	*****		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation.		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>		
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).		
CC	*****		

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
 OC Xenopodidae; Xenopus.  
 OM NCBI\_TaxId=83355;  
 RN [1]  
 RN SEQUENCE OF 23-366 FROM N.A.  
 RP MEDLINE=85074456; PubMed=6210149;  
 RX Ginsberg A.M., King B.O., Roeder R.G.;  
 RA "Xenopus 5S gene transcription factor, TFIIA: characterization of  
 RT cDNA clone and measurement of RNA levels throughout development."; *J*  
 RL Cell 39:479-489(1984).  
 RN [2]  
 RN SEQUENCE OF 23-366 FROM N.A.  
 RP MEDLINE=86176722; PubMed=3754336;  
 RX Yun Tso J., Van den Berg J., Korn L.J.;  
 RA "Structure of the gene for Xenopus transcription factor TFIIA."; *J*  
 RL Nucleic Acids Res. 14:2187-2201(1986).  
 RN [3]  
 RN SEQUENCE FROM N.A., AND SEQUENCE OF 112-137.  
 RP MEDLINE=86312888; PubMed=375818;  
 RX Taylor W., Jackson J.J., Slagel N., Kumar A., Brown D.D.;  
 RA "The developmental expression of the gene for TFIIA in Xenopus  
 RT laevis";  
 RL Nucleic Acids Res. 14:6185-6195(1986).  
 RN [4]  
 RN SEQUENCE OF 1-35 FROM N.A.  
 RP MEDLINE=89306603; PubMed=2744458;  
 RX Scott K.W., Kaulen H., Roeder R.G.;  
 RA "Positive and negative regulation of the gene for transcription  
 RT factor TFIIA in Xenopus laevis oocytes."; *J*  
 RL Genes Dev. 3:651-662(1989).  
 RN [5]  
 RN SEQUENCE OF 1-26 FROM N.A.  
 RP MEDLINE=91071587; PubMed=2253860;  
 RX Kim S.R., Darby M.R., Joho K.E., Brown D.D.;  
 RA "The characterization of the TFIIA synthesized in somatic cells of  
 RT Xenopus laevis";  
 RL Genes Dev. 4:1602-1610(1990).  
 RN [6]  
 RN REPEATS ANALYSIS.  
 RP MEDLINE=85284956; PubMed=4040853;  
 RX Miller J., McLachlan A.D., Kling A.;  
 RA "Repetitive zinc-binding domains in the protein transcription factor  
 RT TFIIA from Xenopus oocytes."; *J*  
 RL EMBO J. 4:1609-1614(1985).  
 RN [7]  
 RN REPEATS ANALYSIS.  
 RP MEDLINE=85231235; PubMed=4007166;  
 RX Brown R.S., Sander C., Argos P.;  
 RA "The primary structure of transcription factor TFIIA has 12  
 RT consecutive repeats";  
 RL FEBS Lett. 186:271-274(1985).  
 RN [8]  
 RN REPEATS ANALYSIS.  
 RP Boehm S., Drescher B.;  
 RA "Multiple internal repeats within the structure of the 5S RNA/DNA  
 RT binding transcription factor TF-TFIIA from Xenopus laevis";  
 RL Studia Biophys. 107:227-247(1985).  
 RN [9]  
 RN STRUCTURE BY NMR OF 10-101.  
 RP MEDLINE=97397344; PubMed=92233405;  
 RX Foster M.P., Wolffe D.S., Radhakrishnan I., Case D.A.,  
 RA Gottesfeld J.M., Wright P.E.;  
 RA "Domain packing and dynamics in the DNA complex of the N-terminal  
 RT zinc fingers of TFIIA";  
 RL Nat. Struct. Biol. 4:605-608(1997).  
 RN [10]  
 RN X-RAY CRYSTALLOGRAPHY (3.1 ANGSTROMS) OF 32-210.  
 RP MEDLINE=98169475; PubMed=9501194;  
 RX Nohle R.T., Conlin R.M., Harrison S.C., Brown R.S.;  
 RA "Differing roles for zinc fingers in DNA recognition: structure of a  
 RT six-finger transcription factor TFIIC complex";  
 RL Proc. Natl. Acad. Sci. 95:1296-1301(1998).

Tue Feb 11 13:12:50 2003

us-09-831-426c-3.rge

Page 1

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 12:48:16 : Search time 3381.18 Seconds

(Without alignments)  
10957.076 Million cell updates/sec

Title: US-09-831-426C-3

Perfect score: 1273

Sequence: 1 atgcgcagcagcagcgcgcga.....cagttactacctgtgctaa 1273

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Geneml:

1: gb\_ba:\*  
2: gb\_hcg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_cm:\*  
21: em\_ov:\*  
22: em\_or:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_pi:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*  
29: em\_vl:\*  
30: em\_hcg\_hum:\*  
31: em\_hcg\_inv:\*  
32: em\_hcg\_other:\*  
33: em\_hcg\_mus:\*  
34: em\_hcg\_pln:\*  
35: em\_hcg\_tod:\*  
36: em\_hcg\_mam:\*  
37: em\_hcg\_vrt:\*  
38: em\_sy:\*  
39: em\_hcgo\_hum:\*  
40: em\_hcgo\_mus:\*  
41: em\_hcgo\_other:\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1233.2	96.9	1381	9 HUMITIA	D32257 Human GTF3A
2	1231.6	96.7	1199	6 E10962	E10962 CDNA's enco
3	1088.8	85.5	1173	9 HSU020272	U20272 Human DNA/R
4	919.2	72.2	1098	9 HSU14134	U14134 Human trens
5	889.8	69.9	1920	9 AK057993	AK057993 Homo sapi
6	719.2	56.5	1298	10 AF391799	AF391799 Mus muscu
7	691	54.3	1071	10 AF391798	AF391798 Rattus no
8	560.6	44.0	996	10 BC004738	BC004738 Mus muscu
9	470.2	36.9	97979	9 AC004739	AC004739 Homo sapi
10	470.2	36.9	193126	9 AC006045	AC006045 Homo sapi
11	452.6	35.6	111722	9 AC074390	AC074390 Homo sapi
12	452.6	35.6	208622	2 AC016444	AC016444 Homo sapi
13	447	35.1	199517	9 AL353741	AL353741 Human DNA
14	418.4	32.9	439	6 AX396111	AX396111 Sequence
15	416.8	32.7	439	6 AX396474	AX396474 Sequence
16	385.4	30.3	163421	2 AC012664	AC012664 Homo sapi
17	385.4	29.5	197709	9 AC073641	AC073641 Homo sapi
18	376	29.5	166141	9 AC012460	AC012460 Human DNA
19	373.4	28.3	200885	2 AC012460	AC012460 Homo sapi
20	372.4	28.3	1693	9 AF265440	AF265440 Homo sapi
21	361.4	28.4	1518	5 XE171114	XE171114 Homo sapi
22	355	27.9	1377	5 XE171114	XE171114 Homo sapi
23	350.8	27.6	1331	5 XE171114	XE171114 Homo sapi
24	340.6	26.8	1297	5 XE171114	XE171114 Homo sapi
25	340.6	26.8	1305	5 XE171114	XE171114 Homo sapi
26	319.2	25.1	12905	9 AC116311	AC116311 Homo sapi
27	319.2	25.1	157165	2 AC104112	AC104112 Homo sapi
28	319.2	25.1	171007	2 AC021381	AC021381 Homo sapi
29	319.2	25.1	171705	2 AC113377	AC113377 Homo sapi
30	284.6	22.4	1314	5 BAF17556	BAF17556 Mus muscu
31	278.6	21.9	67190	2 AC117556	AC117556 Mus muscu
32	234	18.4	234	6 AX396395	AX396395 Sequence
33	221.6	17.4	1129	5 IPOOTF3A	IPOOTF3A Rattus no
34	202	15.9	162774	2 AC113181	AC113181 Rattus no
35	202	15.9	18326	2 AC113563	AC113563 Rattus no
36	197.6	15.5	112902	9 AC005230	AC005230 Homo sapi
37	177	12.9	272	6 AX322436	AX322436 Sequence
38	163	12.8	167342	2 AC124828	AC124828 Mus muscu
39	152.4	12.0	660	6 AX202580	AX202580 Sequence
40	152.4	12.0	660	6 AX202580	AX202580 Sequence
41	152.4	12.0	947	6 AX053174	AX053174 Sequence
42	152.4	12.0	947	6 AX059770	AX059770 Sequence
43	152.4	12.0	947	6 AX201763	AX201763 Sequence
44	152.4	12.0	947	6 AX201936	AX201936 Sequence
45	152.4	12.0	995	6 AX053173	AX053173 Sequence

# ALIGNMENTS

RESULT 1  
LOCUS HUMITIA  
DEFINITION human GTF3A mRNA for xenopus transcription factor IIA homologue,  
complete cds.  
ACCESSION D32257.1 GI:1000446  
VERSION D32257.1  
KEYWORDS GTF3A; Xenopus transcription factor IIA homologue.  
SOURCE Homo sapiens cDNA to mRNA, clone, lib:library of T.Fujiwara, S.Shin  
and Y.Nakamura clone:39H11.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
REFERENCE 1 (bases 1 to 1381)  
AUTHORS Arikawa,H., Nagase,H., Hayashi,N., Ogawa,M., Nagata,M.,

Pred. No. is the number of results predicted by chance to have a

TITLE	Fujiwara, T., Takahashi, E., Shin, S., and Nakamura, Y.
JOURNAL	Molecular cloning, characterization, and chromosomal mapping of a novel human gene ( <i>GTF3A</i> ) that is highly homologous to Xenopus transcription factor IIIA
REFERENCE	Cytogenet. Cell Genet. 70 (3-4), 235-238 (1995)
AUTHORS	95309028 2 (bases 1 to 1381) Nakamura, Y.
TITLE	Direct Submission
JOURNAL	Submitted (22-JUL-1994) Yusuke Nakamura, Cancer Institute,

Department of Biochemistry; 1-3-1 Kami-Iseokuro, Toshima-Ku,  
Tokyo 170, Japan (E-mail: nakamuraeganvx1.jfcr.or.jp,  
Tel: 03-3918-0111 (ex. 4501), Fax: 03-3918-0342)

```
FEATURES
  source      Location/Qualifiers
1. 1381      /organism="Homo sapiens"
           /db_xref="taxon:9606"
```

```
gene      /clone.lib=libarray of T.Fujisawa, S.Shin and Y.Nakamura
          1..1381
          /gene="GTF3A"
CDS       20..1291
          .....
```

/product\_?Xenopus transcription factor I11A homologue  
 /protein\_id=BA006988.1"  
 /db\_xref=GI:161942."  
 /translation="RRSGADAGRCITVARGSVASRDESGASGRGAFAP  
 RGSAGGGLGAGALADPPVVAIESSTLTADPTIAGSSAPTPPALPDRK  
 PDSANSTKAMKLDMLCHPTGERPFCDVGGKGFIDHLSHLITRGE  
 LAAKQKPTKSTNLKMRERHNEQNOQIJCSEDECKTFRKQDLITGCG  
 LFKTQSCGRKRPDSLKLRHAKRASEYCGQSGPANKTVELLNIRTRR  
 CEVARTKFRKDI.LQHKMTAPRDEYCRPGSGRCPTVTEINLSHISE  
 FVCHACCGCTFAMKQSLTRHVAVHPDKMLVKKRSREKREFGLSSQIY  
 QGGLSLCQNGESPNCVEDEKMLSTVAALTKL"  
 386 a 350 c 354 g 291 t  
 BASE COUNT  
 ORIGIN

Query Match	96.98%	Score 1233.2;	DB 9;	Length 1381;
Best Local Similarity	99.58%	Pred. NO. 4.1e-259;		
Matches 1268; Conservative	0;	Mismatches 3;	Indels 3;	Gaps 3;

[illegible]

Oy	481	TGCAGCAGTCGCTGTGATCAAAAAATTCAACGCAAAATCAAACCTTAACAAAAATTGGA	540
Db	499	TGCAGCAGTCGCTGTGATCAAAAAATTCAACGCAAAATCAAACCTTAACAAAAATTGGA	558
Oy	541	ACCCAAACATGAAATATCAAAAAACAATATATATGCAAGTTTGAAGACTGTGAAGAC	600
Db	559	ACCCAAACATGAAATATCAACAAACAAATATATATGCAAGTTTGAAGACTGTGAAGAC	618
Oy	601	CTTTAAGAAACATCAGCAGCTGAAAAATCCATCATGTCGACATACCAATGAACCTCTATT	660
Db	619	CTTTAAGAAACATCAGCAGCTGAAAAATCCATCATGTCGACATACCAATGAACCTCTATT	678
Oy	661	CAAGTGTACCAGGAAGATGTGGGAACACTTTTGATCATCCAGCAAGCTGAACGACA	720
Db	679	CAAGTGTACCAGGAAGATGTGGGAACACTTTTGATCATCCAGCAAGCTGAACGACA	738
Oy	721	TGCGAAGGCCACGAGGGCTATGTATGTCAAAAAGATGTTCTCTTTGTGCAAAAACATG	780
Db	739	TGCGAAGGCCACGAGGGCTATGTATGTCAAAAAGATGTTCTCTTTGTGCAAAAACATG	798
Oy	781	GAGCGAAGCTTGTGAACATGTGAGAGAAACCCATTAAGAGAAATACTATGTGAAGTATG	840
Db	799	GAGCGAAGCTTGTGAACATGTGAGAGAAACCCATTAAGAGAAATACTATGTGAAGTATG	858
Oy	841	CGCGAAACATTTAAAGCAAAAGATTACCTTAAGCAACATCTAAGAACTCATGCCCGCA	900
Db	859	CGCGAAACATTTAAAGCAAAAGATTACCTTAAGCAACATCTAAGAACTCATGCCCGCA	918
Oy	901	AAGGATGTATGTGCTGTGTGCAAGACAGACGCTGTGGAAGAACTATACTACTGTGTTAA	960
Db	919	AAGGATGTATGTGCTGTGTGCAAGACAGACGCTGTGGAAGAACTATACTACTGTGTTAA	978
Oy	961	TCTCCAAAGCCATATCTGTCTCTTCCATGAGAAAGCCGCCCTTTTGTGTGAACATGC	1020
Db	979	TCTCCAAAGCCATATCTGTCTCTTCCATGAGAAAGCCGCCCTTTTGTGTGAACATGC	1038
Oy	1021	TGGCTGTGGCAAAACATTTGCAATGAAACAAAGTCTCAAGGCATGCTGTTCATATGA	1080
Db	1039	TGGCTGTGGCAAAACATTTGCAATGAAACAAAGTCTCAAGGCATGCTGTTCATATGA	1098
Oy	1081	TGCTGACAGAAGAAATGAAGCTCAAAATCTCAAAAAATCTCGTGAAGAAAC-GCACTTTGG	1139
Db	1099	TGCTGACAGAAGAAATGAAGCTCAAAATCTCAAAAAATCTCGTGAAGAAACGGGACCTTTGG	1158
Oy	1140	CGCTCATATCTAGTGGATATATCCCTCTCCCAAAAGAAACAAAGGCAAGCTTATCTTTGT	1199
Db	1159	CGCTCATATCTAGTGGATATAT-CTCTCCAAAGAAACAAAGGCAAGCTTATCTTTGT	1217
Oy	1200	GTCAAAACGAGATGACCCCACTGTGTGGAAACAAAGATGCTCTGACATGTGCATGACAC	1259
Db	1218	GTCAAAACGAGATGACCCCACTGTGTGGAAACAAAGATGCTCTGACATGTGCATGAC	1277
Oy	1260	TTACACCTTGCGTAA 1273	
Db	1278	TTACACCTTGCGTAA 1291	

RESULT 2				
E10962				
LOCUS	E10962	1399 bp	RNA	Linear
DEFINITION	CDNA's encoding hnf1beta protein.			
ACCESSION	E10962.1			
VERSION	E10962.1	GI:22029822		
KEYWORDS	JP 1996070870-A/1.			
SOURCE	Homo sapiens.			
ORGANISM	Homo sapiens.			
REFERENCE	Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.			
AUTHORS	Fujisawa,T., Takeda,S., Shimada,Y., Ozaki,K. and Shin,T.			
TITLE	HUMAN TRANSCRIPTIONAL FACTOR HNF1B GENE			
JOURNAL	Patent: JP 1996070870-A 1 19-MAR-1996:			

Tue Feb 11 13:12:51 2003

us-09-831-426c-4.rge

Page 1

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 10, 2003, 12:48:16 : Search time 3221.82 Seconds

(without alignments)  
10957.076 Million cell updates/sec

Title: US-09-831-426C-4

Perfect score: 1213

Sequence: 1 gtgcgcgcgcgcgcgcgaag.....cagtaactacccttgactaa 1213

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: GenB1.\*  
2: gb\_ba.\*  
3: gb\_brg.\*  
4: gb\_in.\*  
5: gb\_om.\*  
6: gb\_ov.\*  
7: gb\_ph.\*  
8: gb\_pl.\*  
9: gb\_pr.\*  
10: gb\_ro.\*  
11: gb\_sts.\*  
12: gb\_sy.\*  
13: gb\_un.\*  
14: gb\_vl.\*  
15: em\_ba.\*  
16: em\_fun.\*  
17: em\_hum.\*  
18: em\_in.\*  
19: em\_mu.\*  
20: em\_om.\*  
21: em\_or.\*  
22: em\_ov.\*  
23: em\_pat.\*  
24: em\_ph.\*  
25: em\_pl.\*  
26: em\_ro.\*  
27: em\_sts.\*  
28: em\_un.\*  
29: em\_vl.\*  
30: em\_hgt\_hum.\*  
31: em\_hgt\_inv.\*  
32: em\_hgt\_other.\*  
33: em\_hgt\_mus.\*  
34: em\_hgt\_pln.\*  
35: em\_hgt\_rtd.\*  
36: em\_hgt\_mam.\*  
37: em\_hgt\_vrt.\*  
38: em\_sy.\*  
39: em\_hgtg\_hum.\*  
40: em\_hgtg\_mus.\*  
41: em\_hgtg\_other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1171.6	96.6	1381	6 HUMITTA	D32257 Human GF3A
2	1170	96.5	1399	6 HUMITTA	E10962 CDNA's ENCO
3	1088.8	89.8	1173	9 HSDU0272	U20272 Human DNA/R
4	919.2	75.8	1098	9 HSDU4134	U4134 Human Trans
5	889.8	73.4	1920	9 AK057933	AK057933 Homo Sap1
6	719.2	59.3	1298	10 AF391799	AF391799 Mus muscu
7	591	57.0	1071	10 AF391798	AF391798 Rattus no
8	560.6	46.2	996	10 BC032292	BC032292 Homo sap1
9	470.2	38.8	97979	9 AC004739	AC004739 Homo sap1
10	470.2	38.8	193126	9 AC006045	AC006045 Homo sap1
11	452.6	37.3	111722	9 AC074390	AC074390 Homo sap1
12	452.6	37.3	206622	2 AC016444	AC016444 Homo sap1
13	447	36.9	199517	9 AL353741	AL353741 Human DNA
14	418.4	34.5	439	6 AX396111	AX396111 Sequence
15	416.8	34.4	439	6 AX396474	AX396474 Homo sap1
16	385.4	31.8	163421	2 AC012664	AC012664 Homo sap1
17	385.4	31.8	197709	9 AC073641	AC073641 Homo sap1
18	372.4	30.7	200885	2 AC012460	AC012460 Homo sap1
19	372.4	30.7	1693	2 AF263460	AF263460 Homo sap1
20	365.4	29.8	1518	5 XE1FF3A	XE1FF3A Homo sap1
21	355	29.3	1377	5 XE1FF3A	XE1FF3A Homo sap1
22	350.8	28.9	1331	5 XE1FF3A	XE1FF3A Homo sap1
23	340.6	28.1	1297	5 XE1FF3A	XE1FF3A Homo sap1
24	340.6	28.1	1305	5 XE1FF3A	XE1FF3A Homo sap1
25	319.2	26.3	123905	9 AC116311	AC116311 Homo sap1
26	319.2	26.3	157165	2 AC104112	AC104112 Homo sap1
27	319.2	26.3	171007	2 AC013381	AC013381 Homo sap1
28	319.2	26.3	171007	2 AC113377	AC113377 Homo sap1
29	314.4	25.9	16141	5 AL337059	AL337059 Human DNA
30	284.6	23.5	1314	5 BAFPI1A	BAFPI1A Homo sap1
31	278.6	23.0	67190	2 AC117566	AC117566 Mus muscu
32	234	19.3	1234	6 AX396395	AX396395 Sequence
33	221.6	18.3	1129	5 IPOOF3A	IPOOF3A Homo sap1
34	202	16.7	162774	2 AC121381	AC121381 Rattus no
35	197.6	16.3	112902	9 AC005230	AC005230 Homo sap1
36	197.6	14.6	272	6 AX322436	AX322436 Sequence
37	177	13.4	167342	2 AC124828	AC124828 Mus muscu
38	163	13.4	217225	10 AL513345	AL513345 Mouse DNA
39	152.4	12.6	660	6 AX202580	AX202580 Sequence
40	152.4	12.6	947	6 AX053174	AX053174 Sequence
41	152.4	12.6	947	6 AX059770	AX059770 Sequence
42	152.4	12.6	947	6 AX201763	AX201763 Sequence
43	152.4	12.6	947	6 AX201936	AX201936 Sequence
44	152.4	12.6	995	6 AX053173	AX053173 Sequence
45	152.4	12.6	995	6 AX053173	AX053173 Sequence

#### ALIGNMENTS

RESULT 1  
LOCUS HUMITTA 1381 bp mRNA linear PRI 07-FEB-1999  
DEFINITION Human GF3A mRNA for Xenopus transcription factor IIIA homologue,  
complete cds.  
ACCESSION D32257  
VERSION D32257.1 GI:1000446  
KEYWORDS GF3A; Xenopus transcription factor IIIA homologue.  
SOURCE Homo sapiens cDNA to mRNA, clone, library of T. Fujiwara, S. Shin  
and Y. Nakamura Clone:39H11.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
REFERENCE 1 (bases 1 to 1381)  
Atakawa, H., Nagase, H., Hayashi, N., Ogawa, M., Nagata, M.,

Tue Feb 11 13:12:51 2003

us-09-831-426c-4.rge

Page 2

**TITLE** Fujitawa, T., Takahashi, E., Shin, S., and Nakamura, Y. Molecular cloning, characterization, and chromosomal mapping of a novel human gene (GRF3) that is highly homologous to xenopus transcription factor TIIA

**JOURNAL** Cytogenet. Cell Genet. 70 (3-4), 235-238 (1995)

**MEDLINE** 95309028

**AUTHORS** 2 (bases 1 to 1381)

**TITLE** Nakamura, Y.

**JOURNAL** Direct Submission

**TITLE** Submitted (22-JUL-1994) Yusuke Nakamura, Cancer Institute, Department of Biochemistry, 1-37-1, Kami-Ikebukuro, Toshima-ku, Tokyo 170, Japan (E-mail: nakamura@gs.nyu.ac.jp, Tel: 03-3918-0111 (ex 4501), Fax: 03-3918-0342)

**FEATURES** Location (Chromosome)

**FEATURES**  
**SOURCE**

gene	/clone_id=	library of T.fujisawa, S.Shin and Y.Nakamura
CDS	1..1381	/gene=GF3A*
	20..1291	

BASE COUNT  
ORIGIN

Query Match	96.68;	Score 1171.6;	DB 9;	Length 1381;
Best Local Similarity	99.48;	Score 1171.6;	DB 9;	Length 1381;

Matches 1207; Conservative 0; Mismatches 4; Indels 3; Gaps 3;

QY	1	GTGCGGGGGCGCGGGGAAGGTTCAGAGGAGAGCGTGGGCGCGGCGCGCGTCCCGGC	60
Db	80	GTCCCGGCGCTCCCGGAAGTTCAAGACAGCGCGTGGCGGGGCGCG-CGGTCCCGGC	135
QY	61	ACCTGTCTCGGCAAGCTGGCAAGCGCGGCTGGGCTTGAGGAGGCGCGCGCCCTGGGA	120
Db	139	ACGTTGTCTGGAGACCTGGCAAGCGCGGCGCTGGGCGCTTGAGGAGGCGCGCGCCCTGGGA	198
QY	121	TCCGCGCGCGCGCTGTGCCGCGAGTGGGTGGTGTCTGTGACATCGCGGAGCGCTTCATTGC	180
Db	199	TCCGCGCGCGCGCTGTGGCGCGAGTGGGTGGTGTCTGTGACATCGCGGAGCGCTTCATTGC	258
QY	181	AGCGGCGGAGACCTAGCTTCGAGCGCGCGCGCGCGCGCGCTTCCAGAGGTTCACTG	240
Db	259	AGCGGCGGAGACCTAGCTTCGAGCGCGCGCGCGCGCGCGCTTCCAGAGGTTCACTG	318
QY	241	CTCTCTTCCTGACTGACAGCCGCAATTACAGCAAAAGCTTGAAGCTTGAAGCGGCACTGTG	300
Db	319	CTCTCTTCCTGACTGACAGCCGCAATTACAGCAAAAGCTTGAAGCTTGAAGCGGCACTGTG	378
QY	301	CAAGCAACAGGGGGAGAGACCAATTGTTTGTGACTATGAAGAGGTGTGCAAGGCGCTTCAT	360
Db	379	CAAGCAACAGGGGGAGAGACCAATTGTTTGTGACTATGAAGAGGTGTGCAAGGCGCTTCAT	438
QY	361	CAGGACATACATCTGAGACCGCCGCAATTCGTACTGACACAGAGAGAAAAGCGTTGTTTG	420
Db	439	CAGGACATACATCTGAGACCGCCGCAATTCGTACTGACACAGAGAGAAAAGCGTTGTTTG	498
QY	421	TGAGGCACTGGCTGTGATCAAAAATTCAACACACAAAATCAACTGGAAGAAACATTTGA	480
Db	499	TGAGGCACTGGCTGTGATCAAAAATTCAACACACAAAATCAACTGGAAGAAACATTTGA	558

Qy	481	ACCCAAACATGTAATCAACAAAACAAATATATATATGCAAGTTTGCAGACGTAAACAC	540
Db	559	AGCCAAACATGTAATCAACAAAACAAATATATATATGCAAGTTTGAAGCTGTAACACAC	618
Qy	541	CTTTAAGAAACATCAAGACGCGTGAATTCATAGTCCAGCATACCAATGAAACCTCTATT	600
Db	619	CTTTAAGAAACATCAAGACGCGTGAATTCATAGTCCAGCATACCAATGAAACCTCTATT	678
Qy	601	CAAGGTATACCCAGGAAGGATGTGGGAACACTTTCCATACCCAGCAAGCTGAAGACA	660
Db	679	CAAGGTATACCCAGGAAGGATGTGGGAACACTTTCCATACCCAGCAAGCTGAAGACA	738
Qy	661	TGCCAAGGCCCAAGAGGCTATGTATGTCAAAAAGGATGTTCTTTGTGGCAAAAACATG	720
Db	739	TGCCAAGGCCCAAGAGGCTATGTATGTCAAAAAGGATGTTCTTTGTGGCAAAAACATG	798
Qy	721	GACGGAACCTCGAAACATGTGAGAGAAAACCCATTAAGAGAAATCTATGAGAGTATG	780
Db	799	GACGGAACCTCGAAACATGTGAGAGAAAACCCATTAAGAGAAATCTATGAGAGTATG	858
Qy	781	CGCGAAACATTTAAAGCAAGAAATTACCTTAAGCAACATCGAAAACCTCATGCCCGA	840
Db	859	CGCGAAACATTTAAAGCAAGAAATTACCTTAAGCAACATCGAAAACCTCATGCCCGA	918
Qy	841	AAGGATGTATGTGCGCTGTCCAGAGAGAGAGGCTGTGGAAAGAACTATACATGTGTTAA	900
Db	919	AAGGATGTATGTGCGCTGTCCAGAGAGAGAGGCTGTGGAAAGAACTATACATGTGTTAA	978
Qy	901	TCTCCAAACCCATATCTCTCTCTTCATAGAGAAAGCGCGCTTTGTATGTGAACATGC	960
Db	979	TCTCCAAACCCATATCTCTCTCTTCATAGAGAAAGCGCGCTTTGTATGTGAACATGC	1038
Qy	961	TGCGCTGTGGCAAAACATTTGCAATGAAACAAAGTCTCACTAGGCATGCTGTTGACATGA	1020
Db	1039	TGCGCTGTGGCAAAACATTTGCAATGAAACAAAGTCTCACTAGGCATGCTGTTGACATGA	1098
Qy	1021	TCTGTCAAGAAGAAATGAAAGTCAAAATCAATCAAAAATCTGTGAAAAAC-GAGATTGG	1079
Db	1099	TCTGTCAAGAAGAAATGAAAGTCAAAATCAATCAAAAATCTGTGAAAAACGGAGATTGG	1158
Qy	1080	CGCTTCATCTAGTGGATATATATCCCTCCCAAAAGAAACAAGGGCAAGCTTATCTTTGT	1139
Db	1159	CGCTTCATCTAGTGGATATATATCCCTCCCAAAAGAAACAAGGGCAAGCTTATCTTTGT	1217
Qy	1140	GTCGAAACGGAGATCAACCCCAACTGTGTGTGGAAGACAGATGCTCTCGACAGTTGCATAC	1199
Db	1218	GTCGAAACGGAGATCAACCCCAACTGTGTGTGGAAGACAGATGCTCTCGACAGTTGCATAC	1277
Qy	1200	TTACCCCTTGCGTAA 1213	
Db	1278	TTACCCCTTGCGTAA 1291	

RESULT 2				
E10962				
LOCUS	E10962	1399 bp	RNA	Linear
DEFINITION	CDNAs encoding htfIIIA protein.			PAT 29-SEP-1997
ACCESSION	E10962			
VERSION	E10962.1 GI:22028822			
KEYWORDS	JP 1996070870-A/1.			
SOURCE	Homo sapiens.			
ORGANISM	Homo sapiens.			
REFERENCE	Ekura;Ito;Chordata;Cranialata;Vertebrata;Euteleostomi;			
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	Fujiwara,T., Takeeda,S., Shimada,Y., Ozaki,K. and Shin,T.			
JOURNAL	HUMAN TRANSCRIPTIONAL FACTOR IIIA GENE			
COMMENT	Patent: JP 196070870-A I 19-MAR-1996:			
	OTSUKA PHARMACEUT CO LTD			
	OS Homo sapiens (human)			
	JP 196070870-A/I			
	PD 19-MAR-1996			

DT 07-JUL-1996 (first entry)  
 XX  
 DE Transcription factor-IIIA.  
 XX  
 KW Human; transcription factor-IIIA; hTFIIIA; DNA binding protein;  
 KW ribosome; zinc finger; diagnostic; probe; transcription control;  
 KW antitumour; cancer; therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN <EP704526-A1.  
 XX  
 PD 03-APR-1996.  
 XX  
 PF 05-SEP-1995; 95EP-0113908.  
 XX  
 PR 05-SEP-1994; 94JP-0211022.  
 XX  
 PA (SAKA) OTSUKA PHARM CO LTD.  
 XX  
 PI Fujitawa T, Ozaki K, Shimada Y, Shin S, Takeda S;  
 XX  
 DR WPI: 1996-173033/18  
 XX  
 DR N-PSDB: AMT14037, AMT14038.  
 XX  
 PT Human Transcription Factor III A gene - useful in regulation of  
 PT transcription and for diagnosis and treatment of e.g. cancer related  
 PT diseases  
 XX  
 PS Claim 1: Page 8-10; 17pp; English.  
 XX  
 CC The sequence represents human transcription factor-IIIA (hTFIIIA),  
 CC a DNA binding protein which is necessary for the initiation of 5S  
 CC RNA gene transcription, binding to an internal control region of  
 CC the 5S gene. The protein contains 9 zinc finger domains, which are  
 CC homologous to the C2H2 finger domains of Xenopus TFIIIA, except for  
 CC the 6th finger domain which has only 3 amino acid residues between  
 CC 2 cysteine residues, instead of 5 amino acid residues for Xenopus  
 CC TFIIIA. The protein (optionally in recombinant form) and encoding  
 CC gene may be used in diagnosis, identification or therapy of  
 CC hereditary diseases such as cancer, or other diseases resulting from  
 CC abnormal transcriptional control, and to analyse the mechanisms  
 CC involved in their activity.  
 CC  
 XX  
 SQ Sequence 423 AA:  
 Query Match 96.3%; Score 1933; DB 17; Length 423;  
 Best Local Similarity 96.2%; Pred. No. 5, 5e-153;  
 Matches 351; Conservative 3; Mismatches 11; Indels 0; Gaps 0;  
 0;  
 OY 1 MDPRAVAVSVSLTADAFIAGSSAPPPALPRRTICFPCCSANYSKAKLDAH 60  
 DB 59 LDPRAVAVSVSLTADAFIAGSSAPPPALPRRTICFPCCSANYSKAKLDAH 118  
 OY 61 LCAHTGERPVCVDYGCCKAFIRDYHLSRHILHTGCKPPVCAATGCKDKFTKSNLKRH 120  
 DB 119 LCAHTGERPVCVDYGCCKAFIRDYHLSRHILHTGCKPPVCAATGCKDKFTKSNLKRH 178  
 OY 121 FERKHEHNOQKIVICSFEDCKTFFKHOOLKHOCQHTNEPLFCTQEGCGKHFASPSLTK 180  
 DB 179 FERKHEHNOQKIVICSFEDCKTFFKHOOLKHOCQHTNEPLFCTQEGCGKHFASPSLTK 238  
 OY 181 RHAKAHEGVVCGKGSFPAKATITELKHVRETHKEETLCVCKRTKRKDYLOKHAKTHA 240  
 DB 239 RHAKAHEGVVCGKGSFPAKATITELKHVRETHKEETLCVCKRTKRKDYLOKHAKTHA 298  
 OY 241 PENVDCRCRCCGRTYTTVFNLOSHLISREHSRPFVCHACGCTFPAKOSLTRAAV 300  
 DB 299 PENVDCRCRCCGRTYTTVFNLOSHLISREHSRPFVCHACGCTFPAKOSLTRAAV 358  
 OY 301 HOPDKKMKLKYKSRERSLASHLSGVTIPKRNQOGSLCNGESPACVEDKMLSTVA 360  
 DB 359 HOPDKKMKLKYKSRERSLASHLSGVTIPKRNQOGSLCNGESPACVEDKMLSTVA 418

OY 361 VLTLC 365  
 DB 419 VLTLC 423  
 RESULT 4  
 ID ABP41846 standard; protein; 184 AA.  
 XX  
 AC ABP41846;  
 XX  
 DT 22-AUG-2002 (first entry)  
 XX  
 DE Human ovarian antigen HUSYA18, SEQ ID NO:2978.  
 XX  
 KW Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;  
 KW ovarian cancer; breast cancer; tumour; reproductive system disorder;  
 KW infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;  
 KW PCOS; ovarian cyst; dysmenorrhoea; endocrine disorder; infection;  
 KW inflammatory condition; immune disorder; blood disorder;  
 KW cardiovascular disorder; respiratory disorder; neurological disorder;  
 KW gastrointestinal disorder; urinary system disorder; drug screening;  
 KW gene therapy; chromosome mapping; forensic analysis;  
 KW antibody preparation; cytostatic; immunomodulatory; neuroprotective;  
 KW antiinflammatory; gynaecological; reproductive; chromosome 13q12.3-13.1.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200200677-A1.  
 XX  
 PD 03-JAN-2002.  
 XX  
 PF 07-JUN-2001; 2001KO-US18569.  
 XX  
 PR 07-JUN-2000; 2000US-209467P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 XX  
 PI Blaise CE, Rosen CA;  
 XX  
 DR WPI: 2002-147878/19.  
 XX  
 DR N-PSDB: AB054923.  
 XX  
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,  
 PT useful in the prevention, treatment and diagnosis of cancer (e.g.  
 PT ovarian cancer); immune disorders, cardiovascular disorders and  
 PT neurological diseases -  
 XX  
 PS Claim 11; SEQ ID NO 2978; 2922pp; English.  
 XX  
 CC The invention relates to 2175 novel human ovarian antigens (ABP41054-  
 CC ABP43228) and to cDNAs encoding them (AB054131-AB056305), and also  
 CC encompasses polypeptides 908 identical and polynucleotides 958 identical  
 CC to the sequences of the invention. The invention additionally relates to  
 CC recombinant vectors and host cells comprising human ovarian antigen  
 CC polynucleotides, antipeptides against human ovarian antigens, and the use  
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,  
 CC treating, prognosing or preventing various ovary and/or breast-related  
 CC disorders. Such conditions include ovarian cancer and breast cancer, and  
 CC metastatic tumours of ovarian or breast origin, reproductive system  
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,  
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine  
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic  
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and  
 CC vaginitis), immune disorders (e.g., congenital and acquired  
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),  
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,  
 CC respiratory disorders, neurological disorders, gastrointestinal disorders  
 CC and urinary system disorders. Ovarian antigen polypeptides and  
 CC polynucleotides may also be used in screening for compounds which  
 CC modulate ovarian antigen expression or activity. The polynucleotides may  
 CC further be used for gene therapy, chromosome mapping, in the





Oy 961 TCTCAAGCCATATCTCTCTCCATGAGGAAAGCGCCCTTTTGTGTGCAATGC 1020  
 Db 960 TCTCAAGCCATATCTCTCTCCATGAGGAAAGCGCCCTTTTGTGTGCAATGC 1019  
 Oy 1021 TGGCTGTGGCAAAACATTTGCAATCAAAAGCTCTACTAGGCATGCTGTGTACATGA 1080  
 Db 1020 TGGCTGTGGCAAAACATTTGCAATCAAAAGCTCTACTAGGCATGCTGTGTACATGA 1079  
 Oy 1081 TCTGTGACAGAGAAATGAGCTCAAAAGTCAAAATCTCGTGAANAAC - GGAGTTTGG 1139  
 Db 1080 TCTGTGACAGAGAAATGAGCTCAAAAGTCAAAATCTCGTGAANAACGGAGTTTGG 1139  
 Oy 1140 CTTCTCATCTAGTGGATATATCTCTCCAAAGGAACAGGCGCTTATCTTTGT 1199  
 Db 1140 CTTCTCATCTAGTGGATATAT - CTTCCCAAGGAACAGGCGCTTATCTTTGT 1198  
 Oy 1200 GTCAAAACGGAGAGTCACCAACTGTGTGGAAGACAAGATGCTCTGACAGTTGCAAGTAC 1259  
 Db 1199 GTCAAAACGGAGAGTCACCAACTGTGTGGAAGACAAGATGCTCTGACAGTTGCAAGTAC 1258  
 Oy 1260 TTACCCCTTGGC 1270  
 Db 1259 TTACCCCTTGGC 1269

## RESULT 5

AA15407  
 ID AA15407 standard; DNA: 1213 bp.

AC AA15407;

DT 04-SEP-2000 (first entry)

DE Fragment of DNA encoding a transcription factor designated htflIIA.

KW Human; transcription factor; htflIIA; DNA-binding protein;

KW transcription; ribosomal RNA 5S gene; transcriptional control;

KW cancer; ss.

OS Homo sapiens.

XX MO200028024-A1.

XX 18-MAY-2000.

XX 09-NOV-1999; 99WO-FR02738.

XX 10-NOV-1998; 98FR-0014146.

XX (HMRI ) HOECHST MARION ROUSSEL.

XX Bordon-Pallier F, Rocher C;

XX WPI: 2000-387419/33.

XX New nucleic acid encoding human transcription factor IIIA, useful for treatment and diagnosis of cancer and inherited disease ..

XX Claim 4; Page 42; 49pp; French.

XX The present sequence represents a fragment of a human transcription factor (designated htflIIA) gene. The polypeptide is probably a DNA-binding protein probably involved in initiating transcription of the gene for ribosomal RNA 5S and maintaining the stability of transcription of other control genes. The htflIIA polynucleotides and polypeptides are used to make therapeutic or diagnostic compositions for diseases associated with disorders of transcriptional control, particularly cancer or other inherited diseases. The htflIIA polynucleotide can also be used to detect anomalies in gene transcription, particularly for diagnosis of inherited disease, also for studying diseases involving htflIIA.

XX Sequence 1213 BP; 349 A; 309 C; 302 G; 253 T; 0 other;

Query Match 95.28; Score 1211.4; DB 21; Length 1213;  
 Best Local Similarity 99.9%; Pred. No. 4.5e-306;  
 Matches 1212; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 61 GTGCGGGGTGCGCGAAGTTTCAGCAGGAGCGGTGGCGGGCGCGGTTCCCGGC 120  
 Db 1 GTGCGGGGTGCGCGAAGTTTCAGCAGGAGCGGTGGCGGGCGCGGTTCCCGGC 60  
 Oy 121 AGTGTCTCTGGCACGTGGCAGCGCCCTGGCCCTTGGAGCGCGCGCCCTGGA 180  
 Db 61 AGTGTCTCTGGCACGTGGCAGCGCCCTGGCCCTTGGAGCGCGCGCCCTGGA 120  
 Oy 181 TCCGCGCGCGGTGTCGCCGAGTCGTCGTCCTTGGACATCGCGGCGGTTTCATTGC 240  
 Db 121 TCCGCGCGCGGTGTCGCCGAGTCGTCGTCCTTGGACATCGCGGCGGTTTCATTGC 180  
 Oy 241 AGCGCGCGAGAGCTCAGCTCCGAGCCCGCGCGCGCTTCCAGAGGTTTCATCTG 300  
 Db 181 AGCGCGCGAGAGCTCAGCTCCGAGCCCGCGCGCGCTTCCAGAGGTTTCATCTG 240  
 Oy 301 CTCTCTTCCCTGACTGCGAGCGCAATTACAGCAAGGCTTGAAGCTTGACGCGCACCTGTG 360  
 Db 241 CTCTCTTCCCTGACTGCGAGCGCAATTACAGCAAGGCTTGAAGCTTGACGCGCACCTGTG 300  
 Oy 361 CAAGCACACGGGGAGAGACCATTTTGTGTGACTATGAAGGTTGCGCAAGGCTTCAT 420  
 Db 361 CAAGCACACGGGGAGAGACCATTTTGTGTGACTATGAAGGTTGCGCAAGGCTTCAT 360  
 Oy 421 CAGGACTACCATCTGAGCGCGCCACATTCGACTCACAGAGAGAAAGCGTTGTTTG 480  
 Db 361 CAGGACTACCATCTGAGCGCGCCACATTCGACTCACAGAGAGAAAGCGTTGTTTG 420  
 Oy 481 TCGAGCCACTGGCTGTGATCAAAAATTCACACAAAATCAAACTTGAAGAACAATTTGA 540  
 Db 421 TCGAGCCACTGGCTGTGATCAAAAATTCACACAAAATCAAACTTGAAGAACAATTTGA 480  
 Oy 541 ACGCAACATGAAAATCAACAAAACAATATATATGACGTTTGAAGACTGTGAAGAAGAC 600  
 Db 481 ACGCAACATGAAAATCAACAAAACAATATATATGACGTTTGAAGACTGTGAAGAAGAC 540  
 Oy 601 CTTTAAGAAACATCAGCAGCTGAAAATCCATCAGTCGCAGCATACCAATGAACCTTATT 660  
 Db 541 CTTTAAGAAACATCAGCAGCTGAAAATCCATCAGTCGCAGCATACCAATGAACCTTATT 600  
 Oy 661 CAAGTGTACCCAGGAGATGTGGAAACACTTTGCATCACCAGCAAGCTGAACGACA 720  
 Db 601 CAAGTGTACCCAGGAGATGTGGAAACACTTTGCATCACCAGCAAGCTGAACGACA 660  
 Oy 721 TCCCAAGCCCCAGAGGCTATGTATGTCAAAAGGATGTTCTTTTGGCAAAAACATG 780  
 Db 661 TCCCAAGCCCCAGAGGCTATGTATGTCAAAAGGATGTTCTTTTGGCAAAAACATG 720  
 Oy 781 GACGGAATCTCTGAACATGTGAGAGAAACCCATGAAGAGGAATACTATGTGAAGTATG 840  
 Db 721 GACGGAATCTCTGAACATGTGAGAGAAACCCATGAAGAGGAATACTATGTGAAGTATG 780  
 Oy 841 CGGGAACATTTAAACGCAAGATTACCTTAAGCAACACATGAAAACCTATGCCCCAGA 900  
 Db 781 CGGGAACATTTAAACGCAAGATTACCTTAAGCAACACATGAAAACCTATGCCCCAGA 840  
 Oy 901 AAGGATCTATGTGCGCTGTCCAAAGAGAGCGCTGTGAAGAACCTATACTACTGTGTTAA 960  
 Db 841 AAGGATCTATGTGCGCTGTCCAAAGAGAGCGCTGTGAAGAACCTATACTACTGTGTTAA 900  
 Oy 961 TCTCCAAAGCCATATCTCTCTCTTCATGAGGAAGCGCCCTTTTGTGTGAACATGC 1020  
 Db 901 TCTCCAAAGCCATATCTCTCTCTTCATGAGGAAGCGCCCTTTTGTGTGAACATGC 960  
 Oy 1021 TGGCTGTGGCAAAACATTTGCAATCAAAAGCTCTACTAGGCATGCTGTGTACATGA 1080  
 Db 961 TGGCTGTGGCAAAACATTTGCAATCAAAAGCTCTACTAGGCATGCTGTGTACATGA 1020

FH Key Location/Qualifiers  
 FT CDS 20..1291  
 FT /tag= a  
 FT /product= Transcription factor-IIIa  
 FT /note= "Claim 2"  
 FT misc\_feature 317..1096  
 FT /tag= b  
 FT /product= zinc finger domains  
 FT polyA\_signal 1363..1368  
 FT /tag= c  
 FT  
 XX EP704526-A1.  
 XX  
 XX 03-APR-1996.  
 XX  
 XX 05-SEP-1995; 95EP-0113908.  
 XX  
 XX 05-SEP-1994; 94JP-0211022.  
 XX (SAKA ) OTSUKA PHARM CO LTD.  
 XX  
 XX Fujiwara T, Ozaki K, Shimada Y, Shin S, Takeda S;  
 XX  
 XX MPI; 1996-173033/18.  
 XX P-PSDB; AAR91305.  
 XX

Human Transcription Factor III A gene - useful in regulation of transcription and for diagnosis and treatment of e.g. cancer related diseases

Example 1; Page 11-13; 17pp; English.

The sequence encodes human transcription factor-IIIa (hTFIIIA), a DNA binding protein with 9 zinc finger domains, which is necessary for the initiation of 5S RNA gene transcription, binding to an internal control region of the 5S gene. The coding region (claimed) is given in AAT14037. A fragment lacking a 5'-portion of the gene has been isolated from a human foetal brain cDNA library (OTK7-1), and the 5'-portion of the gene has been isolated by 5'-rapid amplification of cDNA ends using primers H11-R (AAT14039), H11-E (AAT14040), H11-H (AAT14041) and AP-2 (AAT14042), complementary to anchor primer AAT14043. Reverse transcription using H-11R is followed by anchor primer ligation and PCR using AP-2 and H11-E, to give a full-length cDNA, OTK7. The gene and its encoded protein may be used in diagnosis, identification or therapy of hereditary diseases such as cancer, or other diseases resulting from abnormal transcriptional control, and to analyse the mechanisms involved in their activity.

Sequence 1399 BP; 405 A; 349 C; 354 G; 291 T; 0 other;

Query Match 95.5%; Score 1170; DB 17; Length 1399;  
 Best Local Similarity 99.3%; Pred. No. 6.5e-303;  
 Matches 1206; Conservative 0; Mismatches 5; Indels 3; Gaps 3;  
 Oy 1 GTCCGCGCGCGCGAGGTTACAGAGGAGCGGTGGCGCGCGCGGTTCGCCGC 60  
 Db 80 GTCCGCGCGGTGCGGGAAGTTACAGAGGAGCGGTGGCGCGCGCGGTTCGCCGC 138  
 Oy 61 ACCTGTCTCGGACGTGGCGCGCGGTGGCGGTGGAGGCGCGCGCGCGGTTCGGA 120  
 Db 139 ACCTGTCTCGGACGTGGCGCGCGGTGGCGGTGGAGGCGCGCGCGGTTCGGA 198  
 Oy 121 TCCGCGCGCGGTGGCGCGCGGTGGCGGTGGAGGCGCGCGCGGTTCGGA 180  
 Db 199 TCCGCGCGCGGTGGCGCGCGGTGGCGGTGGAGGCGCGCGCGGTTCGGA 258  
 Oy 181 AGCGCGGAGAGCTCAGCTCCGACCGCGCGCGGTTCGCGAGGAGGTTCATGTG 240  
 Db 259 AGCGCGGAGAGCTCAGCTCCGACCGCGCGCGGTTCGCGAGGAGGTTCATGTG 318  
 Oy 241 CTCCTTCCCTGACTGCAGCGGCAATACAGCAAGAGCTGGAAGCTTGACGCCGCTGTG 300  
 . . . . .

RESULT 6  
 AAT14037  
 ID AAT14037 standard; cDNA: 1269 BP.  
 XX  
 AC AAT14037;  
 XX

Db 319 CTCCTTCCCTGACTGCAGCGCCAAATTACAGCAAGAGCTTGAAGCTTGACGCGCACCTGTG 378  
 Oy 301 CAAGCACACGGGGAGAGACCATTTGTTTGTGACTATGAGGGTGTGGCAGGCGCTTCAT 360  
 Db 379 CAAGCACACGGGGAGAGACCATTTGTTTGTGACTATGAGGGTGTGGCAGGCGCTTCAT 438  
 Oy 361 CAGGAGCTACCATCTGAGCGCGCCACATTTCTGACTCACAGAGAGAAAAGCCGTTTGTG 420  
 Db 439 CAGGAGCTACCATCTGAGCGCGCCACATTTCTGACTCACAGAGAGAAAAGCCGTTTGTG 498  
 Oy 421 TCGAGCCACTGGCTGTGATCAAAAATTCAACACAAAATCAAACTTTGAAGAACAATTTGA 480  
 Db 499 TCGAGCCAAATGGCTGTGATCAAAAATTCAACACAAAATCAAACTTTGAAGAACAATTTGA 558  
 Oy 481 AGCAAAACATGAAATCAACAAAAACAATATATATGACAGTTTGAAGACTGTGAAGAAGAC 540  
 Db 559 ACGCAACATGAAATCAACAAAAACAATATATATGACAGTTTGAAGACTGTGAAGAAGAC 618  
 Oy 541 CTTTAAGAAACATCAGCAGCTGAAATCCATCCAGTCCAGCATACCAATGAACCTCTATT 600  
 Db 619 CTTTAAGAAACATCAGCAGATGAAATCCATCCAGTCCAGATACCAATGAACCTCTATT 678  
 Oy 601 CAAGTGTACCCAGGAGGATGTGGAAACACTTTTGCATCACCACAGCAAGCTGAACAGACA 660  
 Db 679 CAAGTGTACCCAGGAGGATGTGGAAACACTTTTGCATCACCACAGCAAGCTGAACAGACA 738  
 Oy 661 TGCCAAAGGCCACAGAGGCTATGTATCTCAAAAGGATGTCTTGTGGCAAAAACATG 720  
 Db 739 TGCCAAAGGCCACAGAGGCTATGTATCTCAAAAGGATGTCTTGTGGCAAAAACATG 798  
 Oy 721 GACGGAACCTCTGAAACATGTGAGAGAAACCCATAAGAGGAATACTATGTGAAGTATG 780  
 Db 799 GACGGAACCTCTGAAACATGTGAGAGAAACCCATAAGAGGAATACTATGTGAAGTATG 858  
 Oy 781 CCGGAAACATTTAAACGCAAGATTACCTTAAAGCAACACATGAAAACCTCATGCCCCAGA 840  
 Db 859 CCGGAAACATTTAAACGCAAGATTACCTTAAAGCAACACATGAAAACCTCATGCCCCAGA 918  
 Oy 841 AAGGATGTATGTGCGCTGTCCAAGAGAGGCTGTGGAAGAAGCTATACTACTGTGTTTAA 900  
 Db 919 AAGGATGTATGTGCGCTGTCCAAGAGAGGCTGTGGAAGAAGCTATACTACTGTGTTTAA 978  
 Oy 901 TCTCCAAAGCCATATCTCTCTCTTCCATGAGAAAGCCGCGCTTTGTGTGGAACATGC 960  
 Db 979 TCTCCAAAGCCATATCTCTCTCTTCCATGAGAAAGCCGCGCTTTGTGTGGAACATGC 1038  
 Oy 961 TGGCTGTGGCAAAACATTTGCAATGAAACAAAGTCTCACTAGGCATGCTGTGTACATGA 1020  
 Db 1039 TGGCTGTGGCAAAACATTTGCAATGAAACAAAGTCTCACTAGGCATGCTGTGTACATGA 1098  
 Oy 1021 TCTGACAAAGAAATAATGAAGCTCAAAAGTCAAAAATCTCGTGAANAAC-GGAGTTGG 1079  
 Db 1099 TCTGACAAAGAAATAATGAAGCTCAAAAGTCAAAAATCTCGTGAANAACGGGAGTTGG 1158  
 Oy 1080 CCTCTCATCTCAGTGGATATATCTCCCAAAAGGAAACAAAGGCAAGGCTTATCTTTGT 1139  
 Db 1159 CCTCTCATCTCAGTGGATATAT-CTTCCCAAAAGGAAACAAAGGCAAGGCTTATCTTTGT 1217  
 Oy 1140 GTCAAAACGGAGACTCACCACTGTGTGGAAGACAGATGCTCTCGACAGTTGCGAGTAC 1199  
 Db 1218 GTCAAAACGGAGACTCACCACTGTGTGGAAGACAGATGCTCTCGACAGTTGCGAGTAC 1277  
 Oy 1200 TTACCCCTTGGCTAA 1213  
 Db 1278 TTACCCCTTGGCTAA 1291

